



TITLE:

# Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences

AUTHOR(S):

Fröhlich, Marlen; Kuchenbuch, Paul; Müller, Gudrun; Fruth, Barbara; Furuichi, Takeshi; Wittig, Roman M.; Pika, Simone

---

CITATION:

Fröhlich, Marlen ...[et al]. Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. Scientific Reports 2016, 6: 25887.

ISSUE DATE:

2016-05-23

URL:

<http://hdl.handle.net/2433/226844>

RIGHT:

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

# SCIENTIFIC REPORTS

OPEN

## Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences

Received: 11 February 2016

Accepted: 22 April 2016

Published: 23 May 2016

Marlen Fröhlich<sup>1</sup>, Paul Kuchenbuch<sup>1</sup>, Gudrun Müller<sup>1</sup>, Barbara Fruth<sup>2,3</sup>, Takeshi Furuichi<sup>4</sup>, Roman M. Wittig<sup>5,6</sup> & Simone Pika<sup>1</sup>

Human language is a fundamentally cooperative enterprise, embodying fast-paced and extended social interactions. It has been suggested that it evolved as part of a larger adaptation of humans' species-unique forms of cooperation. Although our closest living relatives, bonobos and chimpanzees, show general cooperative abilities, their communicative interactions seem to lack the cooperative nature of human conversation. Here, we revisited this claim by conducting the first systematic comparison of communicative interactions in mother-infant dyads living in two different communities of bonobos (*LuiKotale*, DRC; *Wamba*, DRC) and chimpanzees (*Tai'South*, Côte d'Ivoire; *Kanyawara*, Uganda) in the wild. Focusing on the communicative function of joint-travel-initiation, we applied parameters of conversation analysis to gestural exchanges between mothers and infants. Results showed that communicative exchanges in both species resemble cooperative turn-taking sequences in human conversation. While bonobos consistently addressed the recipient via gaze before signal initiation and used so-called overlapping responses, chimpanzees engaged in more extended negotiations, involving frequent response waiting and gestural sequences. Our results thus strengthen the hypothesis that interactional intelligence paved the way to the cooperative endeavour of human language and suggest that social matrices highly impact upon communication styles.

"Language didn't make interactional intelligence possible, it is interactional intelligence that made language possible as a means of communication"<sup>1,p.232</sup>

Human communication is one of the most sophisticated signalling systems in the animal kingdom and has often been used to define what it means to be 'human'<sup>2</sup>. Although there is still an ongoing debate concerning which special principles form the core of this ability, most researchers would agree that it is a fundamentally cooperative enterprise<sup>3,4</sup>. The first step into this collective endeavour can already be observed in early infancy, well before the use of first words, when infants start to engage in turn-taking interactional practices embodying gestures to cooperatively share interest in an activity, event or object with other individuals<sup>5</sup>. One of the predominant theories of language evolution (but see for different theories<sup>6</sup>) thus postulates that, phylogenetically, the first fundamental steps towards human communication were not vocalisations, nor a combination of vocal and gestural signals, but were gestures alone<sup>7</sup>. This hypothesis stirred a considerable amount of research attention concerning general gestural abilities of our two closest living relatives, bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*). The resulting studies showed that both species have multifaceted gestural repertoires, which are used as flexible, intentionally produced communicative strategies in a variety of social contexts<sup>8–10</sup>. Although cooperative abilities have clearly been shown under experimental<sup>11,12</sup> and/or natural conditions (for an overview see<sup>13</sup>) in both *Pan* species, some scholars have claimed that their communicative interactions lack the cooperative nature of human communication<sup>1,2</sup>. However, by combining both analytical questions and prior findings from

<sup>1</sup>Humboldt Research Group, Max Planck Institute for Ornithology, Seewiesen, Germany. <sup>2</sup>Department Biology II, Ludwig-Maximilian University, Munich, Germany. <sup>3</sup>Centre for Research and Conservation/KMDA, Antwerp, Belgium.

<sup>4</sup>Primate Research Institute, Kyoto University, Kyoto, Japan. <sup>5</sup>Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. <sup>6</sup>Tai Chimpanzee Project, Centre Suisse de Recherches Scientifiques, Abidjan, Côte d'Ivoire. Correspondence and requests for materials should be addressed to M.F. (email: [mfroehlich@orn.mpg.de](mailto:mfroehlich@orn.mpg.de)) or S.P. (email: [spika@orn.mpg.de](mailto:spika@orn.mpg.de))

comparative research with a conversation analysis framework (termed ‘CA-assisted comparative research’<sup>14</sup>), Rossano<sup>15</sup> recently showed that the underlying structure of bonobo gesturing might be more similar to human conversation, and thus language, than previously thought. Conversation analysts have been intrigued with the question of how social actions can be made intelligible, since intelligibility is required to achieve mutual understanding and facilitates the successful engagement of cooperative interactions. One useful tool for addressing this question is outlining the sequential organization of social action via turns at talk<sup>16</sup>. The most fundamental structure in this organization is the *adjacency pair*, which can be recursively produced and extended in conversation<sup>16,17</sup>. In its minimal, unexpanded form, an adjacency pair has the following features:

- (a) It is composed of two turns,
- (b) by different participants,
- (c) that are adjacently placed, and
- (d) relatively ordered into first pair parts and second pair parts. First pair parts are actions in first position that initiate some exchange (e.g. a request: “shall we leave now?”), and second pair parts are actions in second position that are responsive to first pair parts (e.g. an answer: “yes, perfect timing”)<sup>17</sup>.

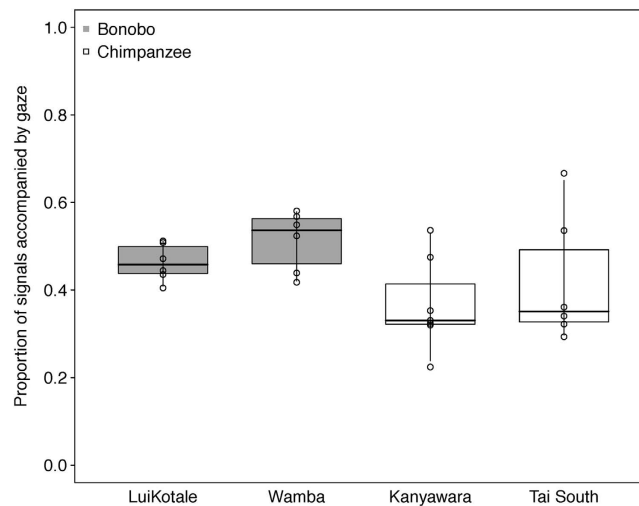
There are clear tendencies for a core inventory of speech acts like questions, greetings, requests, etc. (although many actions beyond the core vary), suggesting a strong universal foundation across all cultures<sup>18</sup>. An additional fundamental part of the infrastructure for human conversation is the temporal relationship underlying turn transitions, which is only about 200 ms on average across languages<sup>19</sup>. This is extraordinary if one bears in mind that latencies involved in uttering even a single word are on the order of 600 ms<sup>20</sup>.

In Rossano’s study, gestural sequences of two mother-infant dyads of bonobos living in captivity were investigated, with a special focus on *participation frameworks* (the signaller decides who is part of the interaction by for instance addressing her visual attention toward a recipient, expecting to turn over the communication role), cooperative *adjacency pair-like sequences* (e.g. request and answer), and temporal relationships underlying gestural performance (e.g. signaller responds immediately after a signal has been produced with a time gap of only  $> 0 < 0.2$  sec)<sup>15</sup>. The results showed that both dyads regularly established and engaged in *participation frameworks* and cooperative *adjacency pair-like sequences* and communicated at a speed remarkably similar to the timing of ordinary human conversation ( $\leq 0.2$  sec).

Although cooperative communication between conspecifics has only been studied in bonobos thus far, other studies provide evidence for communicative differences between the two *Pan* species<sup>10,21</sup>. For instance, Pollick and de Waal<sup>10</sup> found that bonobos in captivity are culturally more diverse in their gesture use and display a higher responsiveness to combinatorial signalling than chimpanzees, giving rise to the speculation that bonobos are a better model for understanding the prerequisites of human communication. Support for this bonobo-chimpanzee dichotomy also stems from other research avenues showing that bonobos are more tolerant and cooperative<sup>11</sup> and outperform chimpanzees in ‘theory of mind’ tasks that require attention to social causality<sup>22</sup>. Recently, these differences in social cognitive abilities and social make-up have been explained not only in relation to brain regions responsible for aversive emotional stimuli eliciting fear and anxiety<sup>23</sup>, but also neural circuitry that may increase empathic sensitivity and prosocial behaviour<sup>24</sup>. The latter findings are especially important in light of shared intentionality<sup>25</sup>, which refers to collaborative interactions in which participants share psychological states with one another<sup>25,26</sup>. Shared intentionality has been suggested as the driving force and “the small psychological difference” in human cognitive evolution that paved the way for the cooperative endeavor of language<sup>26,27</sup>.

The aim of the present study was twofold: First, we revisited the claim that communicative interactions of our closest living relatives, bonobos and chimpanzees, lack the cooperative nature of human communication<sup>1,2</sup>. Second, we investigated whether bonobos are the better model species for understanding the precursors of human communication<sup>10</sup>.

To target these aims, we tested and expanded some of the parameters used by Rossano<sup>15</sup> *in situ* (i.e. in bonobos and chimpanzees living in their natural environments) by focusing on the core niche in which communication is learned—face-to-face interactions of mother-infant dyads. Since it recently had been emphasized that due to relatively high degrees of behavioural plasticity both *Pan* species show considerable inter-site variation (for an overview see<sup>28</sup>), we investigated communicative interactions of two bonobo and two chimpanzee communities at four different sites and locations. In the case of behavioural differences between communities, this study design then offered the possibility to distinguish between within- versus between-species differences<sup>29</sup>. Data were collected at *LuiKotale*, Salonga National Park, Democratic Republic of the Congo (DRC), *Wamba*, Luo Scientific Reserve, DRC, *Kanyawara*, Kibale National Park, Uganda (*Pan troglodytes schweinfurthii*) and *Tai South*, Tai National Park, Côte d’Ivoire (*Pan troglodytes verus*). We focused on the single communicative function of mother-infant joint travel, since previous studies suggested that this context is a promising candidate for the occurrence of frequent turn-taking sequences to achieve a joint goal (leaving a location)<sup>15,30</sup>. The following criteria were analysed: (i) *establishment of participation frameworks* (i.e. the initiator establishes for instance via gaze, body direction who is addressed and will be part of the communicative interaction before performing a gestural signal) by focusing on the parameters of gaze, body orientation and initiation distance; (ii) *adjacency pair-like sequences* (i.e. analysing gestures and their respective responses) by examining the parameters number of gestural requests and their respective responses (so called gesture-response pairs) leading to joint travel and response waiting (i.e. signaller pauses after the signal has been produced for at least two seconds waiting for a response) after each solicitation; and (iii) the *temporal relationships* between joint-travel-initiating behaviour and response, by focusing on the parameters of delayed ( $> 2$  sec), immediate ( $\geq 0 < 2$  sec) and overlapped responses ( $< 0$  sec, for further details see also methods).



**Figure 1. Proportion of gaze before signal initiation in bonobo (grey) and chimpanzee (white) mother-infant dyads as a function of study site.** Dots represent mean proportions per dyad. Indicated are median (horizontal lines), quartiles (boxes) and percentiles (2.5 and 97.5%, vertical lines).

Overall, the results of the present study provide evidence that bonobo and chimpanzee mother-infant dyads frequently engage in cooperative turn-taking sequences, and thereby exhibit many similarities to human social action during conversation. The two species differed significantly with regard to some of the examined parameters: While gaze, close initiation distance and fast-paced responses characterised mother-infant joint travel interactions in bonobos, chimpanzees exhibited a higher number of gesture-response pairs, a higher frequency of response waiting and more delayed responses. By using a combination of methods thus far predominantly applied to human social interactions<sup>16,17</sup>, with an unprecedented within-group, between-group and between species comparison of apes living in natural environments, we show that (a) cooperative communication has a more ancient evolutionary origin than previously thought, and (b) social matrices strongly influence communicative preferences and styles. Our study strengthens the view that human communication represents an ensemble of layers of abilities of different types and different antiquity, with precursor adaptations in turn-taking behaviours in which gestures (and possibly other signals) are embedded<sup>31</sup>.

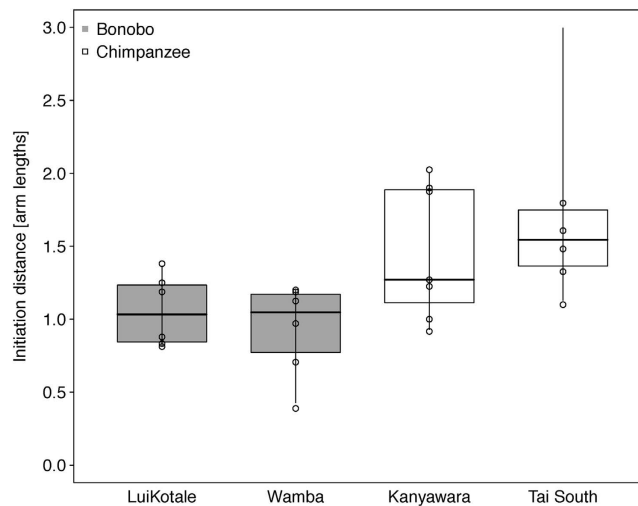
## Results

For our analyses, we distinguished gestures fulfilling only one key characteristic of intentional communication<sup>5,30</sup> (single-criteria or SC-gestures) from those that conformed to several key characteristics of intentional communication (multiple-criteria or MC-gestures). These criteria included parameters such as sensitivity to recipients' attentional states, response waiting and goal persistence<sup>5,30</sup> (for further details see methods and video clips 1–4 in the SM).

The coding of the data set resulted in a total of 400 SC-gestures in bonobos (*LuiKotale*:  $N = 166$ ; *Wamba*:  $N = 234$ ) and 637 in chimpanzees (*Kanyawara*:  $N = 274$ ; *Tai South*:  $N = 363$ ). Concerning MC-gestures, we found a total of 313 gestures in bonobos (*LuiKotale*:  $N = 152$ ; *Wamba*:  $N = 161$ ) and 612 in chimpanzees (*Kanyawara*:  $N = 361$ ; *Tai South*:  $N = 251$ ). For detailed descriptions of the gestures used in joint travel interactions, see Supplementary Material, Table S1.

**Establishment of participation frameworks.** We investigated whether mother-infant dyads established *participation frameworks* before the start of joint travel by analysing the parameters of gaze, body orientation and initiation distance both within and between species. To test the extent to which species and site but also the variables of dyadic role and infant age influenced these three parameters, we used Generalized Linear Mixed Models (GLMM<sup>32</sup>) for each parameter. Overall, the test predictors had a clear impact in all three models (likelihood ratio tests comparing null and the full model for gaze:  $\chi^2 = 16.8$ ,  $df = 8$ ,  $p = 0.03$ ; body orientation:  $\chi^2 = 178.9$ ,  $df = 8$ ,  $p < 0.001$ ; distance:  $\chi^2 = 26.1$ ,  $df = 8$ ,  $p = 0.001$ ). The behavioural differences that we found provided evidence for species differences, but not for within-species variability (likelihood ratio tests comparing null and the reduced model lacking site effects; gaze:  $\chi^2 = 0.2$ ,  $df = 4$ ,  $p = 0.99$ ; body orientation:  $\chi^2 < 0.001$ ,  $df = 4$ ,  $p = 1$ ; distance:  $\chi^2 = 0.009$ ,  $df = 4$ ,  $p = 0.99$ ).

Concerning gaze, species was the only significant predictor, with bonobos accompanying more initiatory behaviours with gaze than did chimpanzees (estimate  $\pm$  SE =  $-0.485 \pm 0.155$ ,  $\chi^2 = 8.288$ ,  $df = 1$ ,  $p = 0.004$ ; see Fig. 1). With respect to body orientation, we found a significant interaction between dyadic role and infant age (role\*between-infants age:  $-0.483 \pm 0.128$ ,  $\chi^2 = 16.307$ ,  $df = 1$ ,  $p < 0.001$ ). While infants oriented more towards the recipient the older they were, there was no effect for mothers with increasing age of their infants. Additionally, chimpanzees more frequently oriented their bodies toward the recipient before signalling than did bonobos ( $0.543 \pm 0.179$ ,  $\chi^2 = 4.872$ ,  $df = 1$ ,  $p = 0.027$ ). Chimpanzee dyads initiated joint travel from a significantly larger distance than did bonobo dyads ( $0.426 \pm 0.103$ ,  $\chi^2 = 7.382$ ,  $df = 1$ ,  $p = 0.007$ ; see Fig. 2). Furthermore, initiation



**Figure 2.** Distance between bonobo (grey boxes) and chimpanzee (white boxes) mother-infant dyads at joint-travel-initiation as a function of study site. Dots represent average initiation distances per dyad. Indicated are median (horizontal lines), quartiles (boxes) and percentiles (2.5 and 97.5%, vertical lines).

distance significantly increased with infant age in both species (between-infants age:  $0.145 \pm 0.041$ ,  $\chi^2 = 9.919$ ,  $df = 1$ ,  $p = 0.002$ ). No other effects in the three models reached significance (see Table 1; sections 1.a-1.c).

**Adjacency pair-like sequences.** To examine whether mother-infant dyads form successful *adjacency pair-like sequences* to initiate joint travel, we investigated two parameters: number of gesture-response pairs and response waiting (i.e. signaller pauses at the end of a given signal for at least two seconds waiting for a response; see also methods). To test the influence of species, site, dyadic role and infant age on these three parameters, we used Generalized Linear Mixed Models (GLMM<sup>32</sup>) for each of them. The test predictors had a clear impact on both models (likelihood ratio tests comparing the null and the full model for gesture-response pairs:  $\chi^2 = 20.9$ ,  $df = 8$ ,  $p = 0.007$ ; response waiting:  $\chi^2 = 42.2$ ,  $df = 8$ ,  $p < 0.001$ ). The differences we found could only be ascribed to inter-species, rather than intra-species, variability (gesture-response pairs:  $\chi^2 < 0.001$ ,  $df = 4$ ,  $p = 1$ ; response waiting:  $\chi^2 < 0.001$ ,  $df = 4$ ,  $p = 1$ ).

Regarding the number of gesture-response pairs, a significant interaction between species and between-infants age was found ( $0.179 \pm 0.071$ ,  $\chi^2 = 6.381$ ,  $df = 1$ ,  $p = 0.012$ ; see Fig. 3). While in bonobos the number of gesture-response pairs decreased, it increased with age in chimpanzees. In addition, mothers produced significantly more gesture-response pairs than infants ( $0.18 \pm 0.078$ ,  $\chi^2 = 5.447$ ,  $df = 1$ ,  $p = 0.02$ ). For the parameter of response waiting, we found a significant interaction between dyadic role and between-infants age ( $-0.433 \pm 0.114$ ,  $\chi^2 = 14.856$ ,  $df = 1$ ,  $p < 0.001$ ). While infants were more likely to wait for a response with increasing age, the age of infants did not influence the occurrence of response waiting in mothers. In addition, chimpanzees were more likely to wait for a response by the recipient than bonobos ( $0.893 \pm 0.173$ ,  $\chi^2 = 19.845$ ,  $df = 1$ ,  $p < 0.001$ ; see Fig. 4). No other effects in the two models reached significance (Tables 1, sections 2.a-2.b).

**Temporal relationships between signal and response.** We examined whether there were inter- or intra-species differences concerning the timing of a given response after a travel-initiating gesture (i.e. SC-gestures and MC-gestures) had been produced. We differentiated between immediate ( $\Delta t$  [start response – end initiatory gesture]  $> 0 < 2$  s), delayed (start response  $\geq 2$  s after end initiatory gesture) and overlapping responses ( $\Delta t$  [start response – end initiatory gesture]  $< 0$ ; for detailed definitions see methods). Bonobos predominantly produced overlapping (N = 200, 47.1%; *LuiKotale*: N = 89, 46.8%; *Wamba*: N = 111, 47.4%) and immediate responses (N = 143, 33.5%; *LuiKotale*: N = 59, 31.1%; *Wamba*: N = 84, 35.9%), but relatively low frequencies of delayed responses (N = 81, 19.4%; *LuiKotale*: N = 42, 22.1%; *Wamba*: N = 39, 16.7%). In contrast, chimpanzees produced approximately equal proportions of overlapping (N = 291, 32.3%; *Kanyawara*: N = 138, 30.5%; *Tai South*: N = 153, 34.1%), immediate (N = 289, 32.1%; *Kanyawara*: N = 157, 34.7%; *Tai South*: N = 132, 29.4%), and delayed responses (N = 321, 35.6%; *Kanyawara*: N = 157, 34.7%; *Tai South*: N = 164, 36.5%). For example see video clips 1–4 in the Supplementary Material. To test the extent to which species, site, dyadic role and infant age influenced the three response types, we used Generalized Linear Mixed Models (GLMM<sup>32</sup>) for each parameter. The test predictors had a clear impact on the occurrence of overlapping and delayed responses, but not on the occurrence of immediate responses (overlapping response:  $\chi^2 = 29.6$ ,  $df = 8$ ,  $p = 0.005$ ; immediate response:  $\chi^2 = 5.6$ ,  $df = 8$ ,  $p = 0.694$ ; delayed response:  $\chi^2 = 35.9$ ,  $df = 8$ ,  $p < 0.001$ ). The likelihood ratio test comparing the full models with the reduced models revealed that behavioural differences mirrored species differences but not within-species variability (overlapping:  $\chi^2 = 0$ ,  $df = 4$ ,  $p = 1$ ; delayed:  $\chi^2 < 0.001$ ,  $df = 4$ ,  $p = 1$ ; see also Supplementary Material, Figure S1).

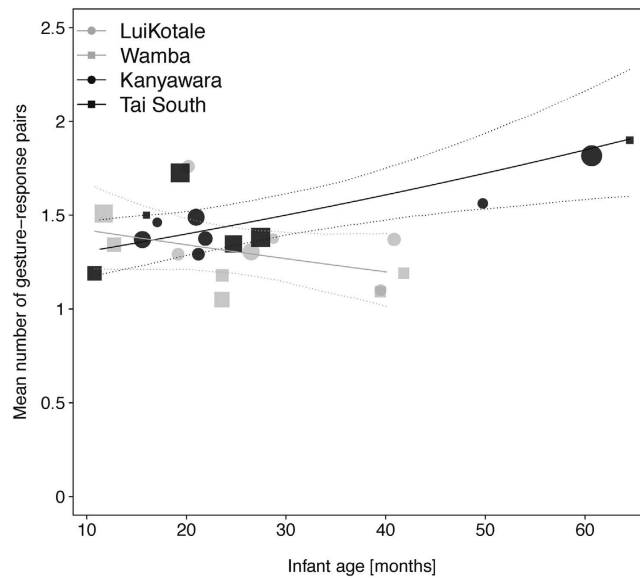
Overlapping responses were significantly more frequent in bonobos ( $-0.65 \pm 0.127$ ,  $\chi^2 = 11.656$ ,  $df = 1$ ,  $p < 0.001$ ; see Fig. 5) than in chimpanzees. Furthermore, infants of both species were more likely to use overlapping responses than were mothers ( $0.382 \pm 0.135$ ,  $\chi^2 = 8.256$ ,  $df = 1$ ,  $p = 0.004$ ). Overlapping responses were

	Estimate	se	$\chi^2$	p
<b>1.a Gaze</b>				
Intercept	−0.121	0.142	(1)	(1)
<b>species [chimp]</b>	<b>−0.485</b>	<b>0.155</b>	<b>8.288</b>	<b>0.004</b>
role [mother]	0.108	0.100	1.157	0.282
within-infants age	−0.053	0.050	1.186	0.276
between-infants age	0.075	0.068	1.152	0.283
infant sex [male]	0.040	0.155	0.068	0.794
parity	−0.009	0.080	0.014	0.906
<b>1.b Body orientation</b>				
Intercept	0.952	0.177	(1)	(1)
<b>species [chimp]</b>	<b>0.543</b>	<b>0.179</b>	<b>4.872</b>	<b>0.027</b>
role [mother]	−1.301	0.117	(1)	(1)
within-infants age	−0.141	0.134	(1)	(1)
between-infants age	0.189	0.142	(1)	(1)
infant sex [male]	0.182	0.186	0.986	0.321
parity	−0.087	0.098	0.753	0.386
role: within-infants age	0.068	0.129	0.275	0.600
<b>role: between-infants age</b>	<b>−0.483</b>	<b>0.128</b>	<b>16.307</b>	<b>&lt;0.001</b>
<b>1.c Initiation distance</b>				
Intercept	0.036	0.097	(1)	(1)
<b>species [chimp]</b>	<b>0.426</b>	<b>0.103</b>	<b>7.382</b>	<b>0.007</b>
role [mother]	−0.095	0.078	1.485	0.223
within-infants age	0.106	0.048	2.923	0.087
<b>between-infants age</b>	<b>0.145</b>	<b>0.041</b>	<b>9.919</b>	<b>0.002</b>
infant sex [male]	−0.080	0.088	0.764	0.382
parity	−0.018	0.046	0.150	0.699
<b>2.a Gesture-response pairs</b>				
Intercept	0.135	0.080	(1)	(1)
species [chimp]	0.129	0.063	(1)	(1)
<b>role [mother]</b>	<b>0.180</b>	<b>0.078</b>	<b>5.447</b>	<b>0.020</b>
within-infants age	−0.038	0.078	(1)	(1)
between-infants age	−0.081	0.065	(1)	(1)
infant sex [male]	−0.052	0.063	0.693	0.405
parity	−0.007	0.030	0.058	0.810
species: within-infants age	0.057	0.084	0.472	0.492
<b>species: between-infants age</b>	<b>0.179</b>	<b>0.071</b>	<b>6.381</b>	<b>0.012</b>
<b>2.b Response waiting</b>				
Intercept	−2.061	0.177	(1)	(1)
<b>species [chimp]</b>	<b>0.893</b>	<b>0.173</b>	<b>19.845</b>	<b>&lt;0.001</b>
role [mother]	0.585	0.137	(1)	(1)
within-infants age	0.154	0.161	(1)	(1)
between-infants age	0.422	0.111	(1)	(1)
infant sex [male]	0.095	0.165	0.327	0.567
parity	0.026	0.082	0.099	0.753
role: within-infants age	−0.234	0.139	2.872	0.090
<b>role: between-infants age</b>	<b>−0.433</b>	<b>0.114</b>	<b>14.856</b>	<b>&lt;0.001</b>
<b>3.a Overlapping response</b>				
Intercept	−0.365	0.142	(1)	(1)
<b>species [chimp]</b>	<b>−0.650</b>	<b>0.127</b>	<b>11.656</b>	<b>&lt;0.001</b>
<b>role [mother]</b>	<b>0.382</b>	<b>0.135</b>	<b>8.256</b>	<b>0.004</b>
within-infants age	−0.091	0.059	1.521	0.217
<b>between-infants age</b>	<b>−0.124</b>	<b>0.063</b>	<b>3.953</b>	<b>0.047</b>
infant sex [male]	−0.032	0.121	0.068	0.794
parity	−0.038	0.062	0.377	0.540
<b>3.b Delayed response</b>				
Intercept	−1.149	0.159	(1)	(1)
species [chimp]	0.917	0.150	(1)	(1)
Continued				

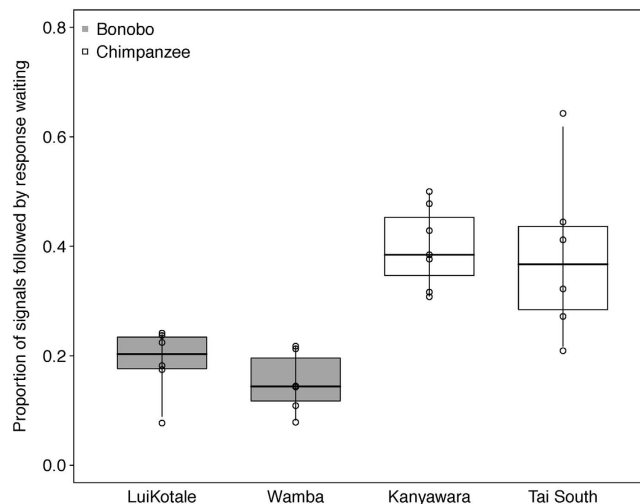


	Estimate	se	$\chi^2$	p
role [mother]	−0.484	0.137	12.416	<0.001
within-infants age	−0.088	0.227	(1)	(1)
between-infants age	0.510	0.177	(1)	(1)
infant sex [male]	−0.004	0.128	0.001	0.975
parity	−0.024	0.063	0.148	0.700
species: within-infants age	0.227	0.242	0.858	0.354
species: between-infants age	−0.510	0.187	7.564	0.006

**Table 1. Effects of species, role and infant age on the investigated parameters.** <sup>(1)</sup>Not shown as lacking a meaningful interpretation. Infant sex and parity were included as control predictors; ID and site were included as random effects.



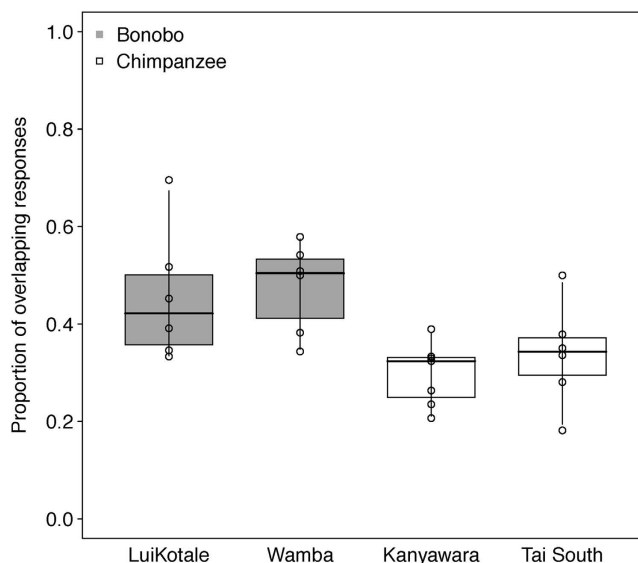
**Figure 3. Count of gesture-response pairs to achieve joint travel in bonobo (grey symbols) and chimpanzee (black symbols) mother-infant dyads of four different communities as a function of infant age.** Depicted are average numbers, separately for each dyad against its mean infant age. The area of the symbols corresponds to the sample size per dyad; the solid and dashed lines represent the fitted model and confidence intervals are based on all other covariates and factors centred to a mean of zero.



**Figure 4. Proportion of signals followed by response waiting in bonobo (grey boxes) and chimpanzee (white boxes) mother-infant dyads as a function of study site.** Dots represent mean proportions per dyad. Indicated are median (horizontal lines), quartiles (boxes) and percentiles (2.5 and 97.5%, vertical lines).

Group	Dyad (infant/mother)	Infant sex	Infant age P1 (months)	Infant age P2 (months)	Interaction time (hours)
LUIKOTALE	Wangila/Wilma	F	14–21	26–28	11.2
	Nora/Nina	F	15–22	28–30	9.1
	Zizu/Zoe	M	22–28	35–37	10.5
	Izzy/Iris	F	26–32	37–40	5.8
	Solea/Susi	F	36–42	48–50	7.3
	Opal/Olga	F	35–42	47–49	8.1
WAMBA	Jolie/Jacky	F	09–13	N/A	8.2
	Seko/Sala	M	10–14	N/A	9.5
	Fua/Fuku	F	20–25	N/A	5.9
	Otoko/Otomi	F	21–25	N/A	6.8
	Hachiro/Hoshi	M	38–42	N/A	5.8
	Kiyota/Kiku	M	39–43	N/A	5.3
KANYAWARA	Winza/Wangari	M	09–11	21–23	15.2
	Tembo/Tenkere	M	13–15	25–27	18.4
	Mango/Michelle	F	13–15	25–27	7.3
	Lily/Leona	F	03–05**	15–17	7.2
	Thatcher/Tongo	F	16–18	28–30	15
	Gola/Outamba	F	48–50*	N/A	7.2
	Wallace/Wilma	M	55–57	67–69	10.1
TAI SOUTH	Mohan/Mbele	F	10–12	22–24	11.2
	Iniesta/Isha	M	N/A	10–12	12
	Solibra/Sumatra	M	15–17	27–29	14.7
	Jeff/Julia	M	15 <sup>†</sup>	N/A	0.4
	Kayo/Kinshasa	F	19–21	31–33	17.0
	Ithaka/Isha	M	64–66*	N/A	9.5
N	25	13:12			238.7

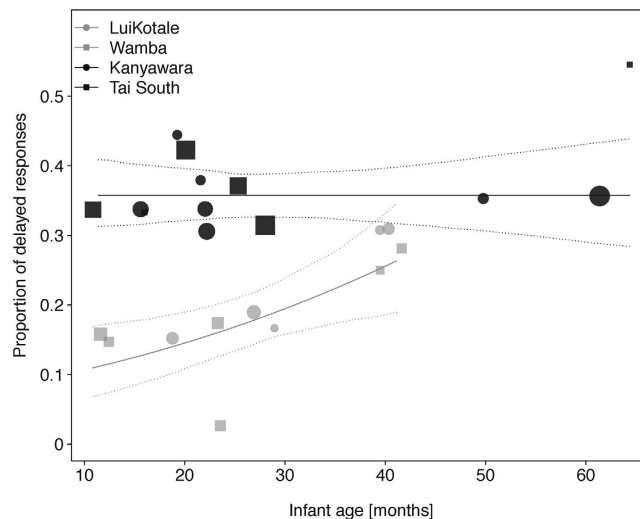
**Table 2. Details on observed dyads as well as respective age/study period, interaction time and sampling efforts.** P1/P2—first/second period of data collection; <sup>†</sup>Deceased on Nov 1, 2012; \*Mothers gave birth to sibling in P2, thus no P2 data available; \*\*P1 not included.



**Figure 5. Proportion of overlapping responses in bonobo (grey boxes) and chimpanzee (white boxes) mother-infant dyads as a function of study site.** Dots represent means per dyad. Indicated are median (horizontal lines), quartiles (boxes) and percentiles (2.5 and 97.5%, vertical lines).

significantly more frequent in dyads with younger infants than in dyads with older infants ( $-0.124 \pm 0.063$ ,  $\chi^2 = 3.953$ ,  $df = 1$ ,  $p = 0.047$ ), irrespective of species. Concerning delayed responses, we found a significant





**Figure 6. Proportion of delayed responses in bonobo (grey symbols) and chimpanzee (black symbols) mother-infant dyads of four different study sites as a function of infant age.** Depicted are proportions, separately for each dyad against its mean infant age. The area of the symbols corresponds to the sample size per dyad; the solid and dashed lines represent the fitted model and confidence intervals are based on all other covariates and factors centred to a mean of zero.

interaction between species and between-infants age, with chimpanzees producing more delayed responses across infant ages than did bonobos ( $-0.51 \pm 0.187$ ,  $\chi^2 = 7.564$ ,  $df = 1$ ,  $p = 0.006$ ; see Fig. 6). Mothers of both species were also more likely to produce delayed responses than their infants ( $-0.484 \pm 0.137$ ,  $\chi^2 = 12.416$ ,  $df = 1$ ,  $p < 0.001$ ). No other effect in the models reached significance (see Table 1, sections 3.a–3.b).

## Discussion

The aim of the present study was twofold: First, we wanted to revisit the claim that human cooperative communication evolved as part of a larger, uniquely human, adaptation for cooperation and cultural life in general<sup>1,2</sup>. Second, we examined whether bonobos are the better model species for understanding the prerequisites of human communication<sup>10</sup>. To do so, we investigated whether bonobos and chimpanzees, both of which engage in general cooperative activities<sup>11–13</sup>, use distinct features characteristic of human social action in conversation<sup>16,17</sup>. By taking into consideration intra- and inter-species variability and by focusing on the mother-infant dyad, our results showed that all observed dyads across groups frequently engaged in turn-taking sequences to negotiate joint travel. They established *participation frameworks* via gaze, body orientation and the adjustment of initiation distance, and they used *adjacency pair-like sequences* characterized by gesture-response pairs and response waiting. Regarding *temporal relationships* between signals and responses, we found that mother-infant dyads of both species used the whole spectrum of responses, including immediate, overlapping and even delayed responses. Immediate responses match the temporal relations between turns in human speech consisting of relatively little cultural variation (e.g. overall cross-linguistic median of 100 ms, ranging from 0 ms in the English and Japanese culture, for instance, to 300 ms in the Danish and Lao culture)<sup>19</sup>. Our findings therefore support and expand the results of Rossano<sup>15</sup>, by demonstrating that gestural exchanges observed in mother-infant dyads of bonobos and chimpanzees are often very similar in timing to human action in conversation and embody the most crucial features of human cooperative conversation<sup>4,27</sup>: These gestural exchanges are bidirectional coordination devices, comprising the two implicit roles of signaller and recipient. In learning to use these gestures, individuals learn to play and to comprehend both roles no matter which role they are performing (soliciting to leave a location or being solicited to leave a location). Although we recently showed that chimpanzee mothers and their infants differ in many of the gesture types employed to initiate joint travel<sup>30</sup>, gestures shared by mothers and infants might involve role-reversal imitation (i.e. when one uses a gesture toward others the way others have used this gesture toward oneself), but also take the other's perspectives on the event of joint travel. Furthermore, to reach the joint goal (of leaving a location), signallers made efforts to communicate in ways that were comprehensible to the recipient, for instance by combining initiatory behaviours with gaze and orienting the body to recipients. In addition, they seemed to 'clarify' the intended goal by using several *adjacency pair-like sequences* composed of the same or different gestures, when the first communicative attempt had not been successful. Turn-taking sequences of pre-linguistic human children go a step further in that recipients ask for clarification when needed and employ 'negotiation of meaning'<sup>33</sup>. Furthermore, the rapid turn-taking in human conversation involves indefinite varying contents of turns, multi-modal deployment of vocal and gestural signals, and also seems without parallel, given the sheer amount of time and effort invested in communication<sup>31</sup> (but see for nonhuman primates and birds<sup>34,35</sup>).

Overall, our findings strengthen a recent proposal by Levinson and Holler<sup>31</sup> emphasizing the role of turn-taking behaviour for evolutionary scenarios of human language. They suggest that human language, despite its tight integration of speech and gesture, is a system composed of layers of abilities of different types and different antiquity. Thus, unpeeling the layers should enable us to understand the evolution of human language from

an original rapid exchange of gestural or vocal material, into a system where the complexity of the linguistic and gestural material that is expressed in relatively short bursts has grown to the very limits that human cognition can process<sup>31</sup>.

In sum, sequentially organized, cooperative social interactions are not simply by-products of individuals living in human enculturated environments<sup>15</sup>, but play a crucial role in communicative exchanges of mother-infant dyads of bonobos and chimpanzees living under active selection pressures. These results challenge the human-ape divide, which suggests that human cooperative communication evolved as part of a larger adaptation of humans' species-unique forms of cooperation<sup>2</sup> ratcheted via existing and simpler components of primate cognition, such as group action and manipulative communication<sup>26</sup>. Our findings indicate that cooperative communicative interactions seem to play a crucial role in mother-infant dyads of bonobos and chimpanzees and, more generally, in nonhuman animals, for which shared goals and relatively low levels of competition prevail. Similarly to the universally organized social-interaction matrix of human conversation<sup>19</sup>, the results suggest that our closest living relatives have a strong universal infrastructure underlying their gestural interactions, which serves to minimize gaps and overlaps and allows for efficient information exchange. Further research on the methods and model species commonly used to draw inferences about evolutionary precursors to human communication is warranted (see for recent developments in other areas of cognitive ethology<sup>36</sup>), to enable (i) higher sensitivity to the social characteristics and/or ecology of a given species and (ii) a vital understanding of the structure and cognitive complexity underlying turn-taking sequences<sup>14</sup> and communicative exchanges such as vocal alternations<sup>34</sup> and duetting<sup>37</sup>.

Concerning our second aim of examining whether bonobos are the better models for precursors to human communication, a comparison of the investigated parameters showed behavioural differences between species, but not within species. Specifically, bonobo dyads (i) accompanied their signals more frequently with gaze, (ii) stayed in closer spatial proximity to each other for mother-infant coordination, and (iii) preferred to use overlapping and immediate responses. In contrast, chimpanzee dyads (i) were more likely to orient their bodies toward a recipient before signalling, (ii) showed a higher number of gesture-response pairs and response waiting, (iii) displayed overall more 'communicative persistence' to obtain the desired goal of joint travel, and (iv) used all three response tempi with relatively similar frequencies.

Three hypotheses may account for these observations. First, differences in communicative patterns may be explained by differences in ecological environments. For example, habitat characteristics such as thickness and growth of terrestrial herbaceous vegetation (THV) may differ considerably between the sites, resulting in different degrees of visibility and thus communication space and eye contact. Although THV is clearly more prevalent in bonobo habitats than in chimpanzee habitats<sup>38</sup>, differences in THV might also exist between the two chimpanzee habitats, resulting in relatively higher levels of visibility at *Kanyawara*<sup>38</sup> compared to *Tai South*<sup>39</sup>. If this hypothesis were true, we would have expected to find differences in communicative behaviours between the *Kanyawara* and the *Tai South* community and between bonobos and chimpanzees in general. This does not accord with our observations.

Second, differences in communication styles between bonobos and chimpanzees are a by-product of the studied age range, particularly because chimpanzee infants may generally develop more quickly than do bonobo infants. For instance, Kuroda<sup>40</sup> suggested that growth rates of bonobos and chimpanzees differ considerably, such that bonobos undergo a slower development of (i) spatial independence, (ii) locomotor skills (e.g. climbing, walking quadrupedally, riding on mother's back), and (iii) social interactions with conspecifics (e.g. approaching, playing). This proposed delay in general development in bonobos may also have a crucial impact on the speed of communicatory skill development. If this hypothesis were true, we would have expected to find that age had a significant impact on the investigated parameters, with bonobo infants showing certain parameters such as use of gaze, adjustment of body orientation and response waiting, as well as overlapping responses, significantly later than chimpanzee infants. However, this was not the case. Body orientation and initiation distance were the only parameters for which a developmental effect was found, with increases of adjustment of body orientation towards mothers and initiation distance with age in both bonobo and chimpanzee infants. Overall, chimpanzee dyads initiated joint travel from larger distances than did bonobo dyads. These results are in line with findings of de Lathouwers and colleagues<sup>41</sup>, who showed that immature chimpanzees spend more time at larger distances from their mothers than do immature bonobos. Our study confirms that chimpanzees indeed develop spatial independence more quickly than do bonobos.

Third, bonobos and chimpanzees might employ different communication styles. Consistent with this hypothesis and based on our investigated parameters, bonobos and chimpanzees could be characterized by two clearly distinguishable communication styles accompanied by different temporal relationships: Bonobos frequently combined their communicative signals with gaze while in close proximity to the addressee, and they often used speedy responses. The underlying temporal relationships often matched those underlying human turn transition during speech<sup>19</sup>, with predominantly single *adjacency pairs* but also recipients responding before signals had been fully articulated. Chimpanzees, on the other hand, adjusted their body orientation toward recipients and used a generally slower mode of communication that involved more gesture-response pairs, higher frequencies of response waiting and delayed responses. Bonobo communication thus seems to resemble a subtle dance coined of flowing movements by signallers and recipients, while chimpanzee communication is structured with temporally separated and clearly recognizable units such as signal, pause and response. Chimpanzee signalling mirrors the structure of other social interactions, such as aggressive and grooming interactions, which are also characterized by typical negotiation sequences<sup>42,43</sup>, thereby demonstrating the significance of clearly structured interactions in chimpanzee society.

While future studies with additional age classes, communicative functions and dyads are of course mandatory, our study provides the first evidence that mother-infant dyads of bonobos and chimpanzees living in their natural environments employ different communication styles to convey the same message. Moreover, if certain

communicative patterns are already observable in mother-infant coordination, the first step of co-regulated social interaction<sup>44</sup>, it is likely that these patterns are also crucial for general communication abilities of the species. Thus, generalization to behaviours of other dyads of a given community may be possible to some extent.

Although the long-standing bonobo-chimpanzee dichotomy has been challenged by new data emphasizing intra-species over inter-species variability<sup>28</sup>, bonobos and chimpanzees still seem to differ considerably concerning distinct characteristics of their social matrices. Males are more influential in chimpanzee society than in bonobo society<sup>45,46</sup>, with male chimpanzees heavily competing within their communities to gain indirect and direct fitness benefits<sup>13,46</sup>. This competition results in linear dominance hierarchies, male harassment and male-female dominance, but also strong social bonds and cooperative behaviour between males in the form of short- and long-term alliances (e.g. in the form of coalitionary behaviour, grooming, meat sharing and border patrols<sup>13,46</sup>). High levels of aggression, including lethal attacks, characterize intergroup encounters in chimpanzees, and infanticide has been observed within and between communities<sup>47</sup>. In contrast, bonobo society is characterized by co-dominance between the sexes, prolonged mother-son relationships<sup>45</sup>, and strong bonds between unrelated females<sup>45,48</sup>, resulting in a more flexible choice of coalition partners. Although between-group encounters in bonobos are usually friendly and peaceful<sup>49</sup>, there is anecdotal evidence of attempts of infanticide by males<sup>50</sup> but also females<sup>51</sup>. Given the gregariousness of bonobo females, the threat of female infanticide could explain the need for close range communication between mothers and their dependant offspring.

It has been argued that species-specific social matrices and behaviours have been evolutionarily shaped by the distinctive morphology, connectivity and molecular biology of brain regions and pathways involved in social and environmental appraisal of threats and vigilance and control of emotional responses<sup>23,24</sup>. For instance, bonobos have more gray matter in the dorsal amygdala and a larger pathway linking the amygdala with the ventromedial prefrontal cortex (VMPFC)<sup>24</sup>. This neural circuitry has been implicated in both top-down control of aggressive impulses and bottom-up biases against harming others, as well as increased empathic sensitivity and prosocial behaviour<sup>24</sup>. In addition, bonobos have approximately twice the density of serotonergic axons in the amygdala compared to chimpanzees<sup>51</sup>, contributing to appraisal of the emotional context and significance of the environment<sup>52</sup>. These differences in neural circuitry are in line with recent experimental findings showing that bonobos tend to exhibit more cautious temperaments<sup>53</sup>, reduced 'emotional reactivity'<sup>11</sup>, and greater tolerance when competing over food-resources<sup>36</sup>. The results of our study suggest that crucial features characterizing human communication, such as gaze and anticipation of recipients' behaviour<sup>2,5</sup>, may be more significant in bonobo than in chimpanzee communication. Bonobos appear to exhibit a higher social awareness of the communicative situation and the anticipated meaning of a given signal, strengthening recent results demonstrating a bonobo-chimpanzee divergence in tasks requiring attention to social causality<sup>22</sup>.

Bonobos may therefore represent the most representative model for understanding the prerequisites of human communication<sup>10</sup>. However, additional analyses of the communicative and cognitive abilities of our closest living relatives are compulsory for a complete understanding of the impact of social and possibly cultural matrices on communication styles and tendencies. In addition, examples of convergent evolution in distantly related species can provide clues to the types of problems that particular communicative mechanisms are 'designed' to solve<sup>54,55</sup>. We thus hope to inspire future research that not only incorporates additional dyads and contexts, but also conducts taxonomically informed comparisons of species engaging in turn-taking behaviour during general interactions and communicative exchanges.

In sum, our results provide substantial evidence that the two primary model species for the origins of human behaviour, bonobos and chimpanzees, differ in their communication styles. While bonobos seem to anticipate and respond to signals before they have been fully articulated, chimpanzees engage in more time-consuming communicative negotiations. Both species use sequentially organized, cooperative social interactions to engage in a joint enterprise: Leaving together to another location. Their communicative interactions thus show the hallmarks of human social action during conversation and suggest that cooperative communication arose as a way of coordinating collaborative activities more efficiently. Our results strengthen a recent proposal by Levinson and Holler<sup>31</sup> suggesting that the apparent gulf between animal and human communication may be bridged by looking for precursors adaptations to human language in turn-taking interactions.

## Methods

**Study sites and subjects.** The study was conducted at two communities of bonobos (*LuiKotale* at the fringe of Salonga National Park, DRC; *Wamba* in the Luo Scientific Reserve, DRC) and two communities of chimpanzees (*Kanyawara* in Kibale National Park, Uganda; *Tai South* in Tai National Park, Côte d'Ivoire). Detailed descriptions of the study areas can be found in Hohmann and Fruth<sup>56</sup> for *LuiKotale*, Kano<sup>57</sup> for *Wamba*, Wrangham and colleagues<sup>58</sup> for *Kanyawara* and Boesch and Boesch-Achermann<sup>39</sup> for *Tai South*. The communicative behaviour of bonobos was observed at *LuiKotale* by P.K. and a trained field assistant from April to November 2012 and from April to July 2013. A second trained field assistant collected data at *Wamba* from September 2012 to February 2013. The behaviour of chimpanzees was observed by M.F. during four study periods between October 2012 and June 2014 (*Kanyawara*: Mar–May 2013, Mar–Jun 2014; *Tai South*: Oct–Dec 2012, Oct–Dec 2013). During the study periods, the number of community members varied between 35 and 40 at *LuiKotale*, around 31 at *Wamba*, between 53 and 56 at *Kanyawara* and between 26 and 33 individuals at the *Tai South* community. Bonobos and chimpanzees at all four sites were well habituated to human observers and have been studied on a longitudinal basis since 2003 in *LuiKotale*<sup>56</sup>, 1973 in *Wamba*<sup>57</sup>, 1987 in *Kanyawara*<sup>58</sup> and 1979 in *Tai South*<sup>39</sup>. It was therefore possible to observe the community members during dawn-till-dusk follows and to collect high-quality video and audio footage. In addition, we had access to long-term data on demography and relatedness for all four field sites. We observed communicative interactions between mothers and their youngest dependent offspring in a total of 12 bonobo dyads and 13 chimpanzee dyads: Six dyads were observed at *LuiKotale*, six at *Wamba*, seven at

*Kanyawara* and six at *Tai South*. The age of the offspring ranged from ten to 56 months in bonobos and nine to 69 months in chimpanzees (see Table 2 for detailed information on subjects and data sets).

**Data collection.** We used a focal behaviour sampling approach<sup>59</sup>, while maintaining a record of the frequency with which a particular dyad had been observed. In situations where we could choose which of several dyads to film, we targeted those individuals previously sampled least often. All social interactions involving mothers and infants (i.e. mother-infant as well as mother-conspecific and infant-conspecific interactions) that were judged to have any potential for communicative interactions were recorded using a digital High-Definition camera (Canon Legria HF M41) with an externally attached unidirectional microphone (Sennheiser K6). During approximately a total of 2200 hours of observation (1033 hours for bonobos; 1189 hours for chimpanzees), we collected a total of 238.5 hours of video footage on the communicative behaviour of 12 bonobo (*Wamba*: 41.5 h; *LuiKotale*: 51.9 h) and 13 chimpanzee (*Tai South*: 73.4 h; *Kanyawara*: 95.5 h) mother-infant dyads (mean  $\pm$  SD per dyad =  $9.5 \pm 3.5$  h; see Table 2 for further details on data collection). However, since this paper focuses only on the communicative context of joint travel, our analysis is based on a total of 319 bonobo and 410 chimpanzee video recordings of mother-infant joint travel interactions (mean recordings per bonobo/chimpanzee dyad: 26.6/31.9; see Supplementary Material, Table S1). In addition, we included five joint-travel interactions that were recorded with a Pocket PC (HP iPAQ rx1959), resulting in a total of 415 interactions for chimpanzees.

**Coding of behaviours.** To establish the repertoires employed to initiate joint travel and to enable subsequent analyses, a total of 729 high-quality video files of mother-offspring joint-travel-initiations (i.e. carries with clear visibility of joint-travel-initiating behaviours) were coded using the program Adobe Premiere Pro CS4 (version 4.2.1). Behavioural definitions were based on established ethograms of bonobo<sup>57</sup> and chimpanzee behaviour<sup>60</sup>. The coding scheme was designed by using parameters developed in previous work on great ape gesturing<sup>30,61</sup>. We only coded successful, agent-initiated joint travel interactions, meaning those interactions leading to infants leaving a location attached to their mothers in a dorsal or ventral carry position. We differentiated between joint-travel-initiations via intentionally produced gestures showing more than one key characteristic of intentional communication and gestures showing only one key characteristic of intentionality, which are respectively called multiple-criteria gestures (MC-gestures) and single-criteria gestures (SC-gestures). Gestures were defined as directed, mechanically ineffective movements of the body or body postures that elicited (“requested”) a voluntary response by the recipient<sup>62</sup>.

The following key characteristics of intentional communication were measured<sup>5,30</sup>:

*Sensitivity to the attentional state of the recipient.* The signaller shows signs of being aware of the recipient’s state of attention, e.g. by using visual gestures only when the recipient is looking.

*Response waiting.* The signaller pauses at the end of the production of a signal and waits for at least two seconds expecting a response, while maintaining visual contact with the recipient.

*Apparent satisfaction of signaller.* The signaller’s communication ceases when the apparent goal has been met by the recipient (leaving the location together).

*Goal persistence.* The signaller elaborates her signalling when thwarted, e.g. by repeating and exaggerating the signal or by using a different means of communication.

**Establishment of participation frameworks.** We examined three different parameters:

*Gaze.* Signaller looks at the recipient while executing the gesture.

*Body orientation.* Recipient is positioned directly in front of and in the visual field of the signaller.

*Initiation distance.* Physical distance (in arm lengths) between signaller and recipient at first joint-travel-initiating gesture.

**Adjacency pair-like sequences.** To investigate whether bonobos and chimpanzees use adjacency pair-like sequences, we focused on gesture-response pairs and response waiting (see paragraph above for definition) after each solicitation gesture. Regarding gesture-response pairs, we analysed the gestures and gestural bouts that included single gestures or sequences, separated by periods of response waiting lasting more than a second and followed by a recipient’s response.

**Temporal relationships between turns.** We assessed the temporal relationships between signals and their respective responses by differentiating between three types of responses (although see for exact times Supplementary Material, Figure S1):

*Immediate response.* The joint-travel-initiating behaviour is followed by a response of the recipient ( $\Delta t$  [start response – end of initiatory gesture]  $> 0$ ) less than two seconds after the behaviour has been articulated.

*Delayed response.* The joint-travel-initiating behaviour is followed by a response of the recipient more than two seconds after the behaviour has been articulated (end of initiatory gesture).



**Overlapping response.** The joint-travel-initiating behaviour (SC- or MC- gesture) is followed by a response of the recipient either before the behaviour has been fully executed ( $\Delta t$  [start response – end initiatory gesture]  $< 0$ ), or less than one second after it has been fully articulated. The full articulation was deciphered by the observer when a given gesture was followed by the immediate reaction (on behalf of the recipient) or response waiting (on behalf of the signaller).

For each gesture used to solicit joint travel, we coded the interaction role of the actor (2 levels: mother, infant), infant age (range = 9–69 months), infant sex (2 levels: female, male), and mother's parity (range = 1–5 offspring). Fifteen per cent of all mother-infant interactions were coded for accuracy by a second observer and tested using the Cohen's Kappa coefficient to ensure inter-observer reliability<sup>59</sup> with the following results: A 'very good' level of agreement for initiatory gesture type/gaze ( $\kappa = 0.80$ ), joint-travel-initiator ( $\kappa = 0.84$ ), orientation and distance to recipient ( $\kappa = 0.89$ ), gesture-response pairs ( $\kappa = 0.89$ ), and temporal relationships including response waiting ( $\kappa = 0.85$ ).

**Analyses.** To test the extent to which species and/or site, but also other parameters such as dyadic role and infant age, influenced i) the *establishment of participation frameworks* before signalling (response variables: gaze, body orientation, initiation distance), ii) successful *adjacency pair-like sequences* (response variables: number of gesture-response pairs, response waiting), and iii) the timing of signal and response (response variables: immediate response, delayed response and overlapping response), we used Generalized Linear Mixed Models (GLMM<sup>32</sup>) with a binomial error structure and logit link function. For the number of gesture-response pairs and distance, we used a Poisson error structure and log link function. We included species, dyadic role and infant age as our key test predictors. Since age varied considerably between infants, we used the method of within-subject centering<sup>63</sup> to determine whether the effect of infant age was particularly relevant within and/or between infants. Hence, we included in the model the average age of each infant (constant across all data points of the respective mother-infant pair; 'between-infants age') and the difference between the infant's actual age and its average age ('within-infants age'). Because we predicted differences between the species and assumed that infants would take a more active role throughout ontogeny, we also included four two-way interactions between both species and role with the two variables representing infant age in the model. To control for the effect, we also included infant's sex and mother's parity as fixed effects in the model. We included study site and identity of the mother and infant as random effects (intercepts). To keep type 1 error rates at the nominal level of 5%, we also included role, within-infants age, parity and infant sex within subject identity, and site<sup>64</sup> as random slopes components. We did not include any other random slopes components within mother ID because, with a single exception, each mother was recorded with only a single infant, so the random slopes of these fixed effects within mother ID would be highly redundant with those within infant identity. For the other fixed effects, we did not include random slopes since they were usually constant within mother and infant ID. To keep model complexity at an acceptable level, and because neglected random slopes do not compromise type 1 error rates<sup>64</sup>, we did not include correlations between random slopes and random intercepts.

The models were implemented in R<sup>65</sup> using the function *glmer* in the package *lme4*<sup>66</sup>. To test the overall significance of our key test predictors, we used a likelihood ratio test<sup>67</sup> to compare the full models with a null model that contained only the control predictor with fixed effects and all random effects. If either interaction with within- and between-infants age was non-significant, they were removed from the model. To test whether inter-site differences had a significant effect on the response variables, we excluded site (and all random slopes within site), and ran a second likelihood ratio test comparing the full model with this reduced model. Prior to running the models, we z-transformed between-infants age, within-infants age and parity<sup>68</sup>. To control for collinearity, we determined Variance Inflation Factors (VIF<sup>69</sup>) from a model that included only the fixed main effects, using the function *vif* of the R package *car*. This revealed that collinearity was not an issue (maximum VIF = 1.23). In the models with Poisson error structure, overdispersion was not an issue (dispersion parameters for distance: 1.03, gesture-response pairs: 0.52). To estimate model stability, we excluded the levels of random effects one at a time, ran the models again and compared the resulting estimates derived with those obtained from the respective models based on all data. This revealed that all models were at least 'moderately' stable, particularly for those estimates that were not close to zero. Tests of the individual fixed effects were derived using likelihood ratio tests (R function *drop1* with argument 'test' set to "Chisq"). All statistical analyses were performed using R-version R.3.1.1<sup>65</sup>, with the level of significance set to 0.05.

**Ethics statement.** Our study was purely non-invasive, with audio and video recordings taken from a minimum distance of seven meters, in an effort to avoid influencing the natural behaviour of the individuals, parties and communities. The research adhered to the legal requirements of the countries in which it was conducted and followed the recommendations of the 'Animals (Scientific Procedures) Act 1986', as published by the government of the United Kingdom, and the principles of "Ethical Treatment of Non-Human Primates" as stated by the American Society of Primatologists. Permission to conduct research at the four field sites was granted by the Centre de Recherche en Écologie et Foresterie (CREF; DRC), the Institut Congolaise pour la Conservation de la Nature (ICCN; DRC), the Makerere University Biological Field Station (MUBFS; Uganda), the Ministère de l'Enseignement Supérieure et de la Recherche Scientifique (Côte d'Ivoire), the Ministère de Recherche Scientifique (DRC), the Office Ivoirien des Parcs et Réserves (OIPR; Côte d'Ivoire), the Uganda National Council for Science and Technology (UNCST; Uganda) and the Uganda Wildlife Authority (UWA; Uganda).

## References

1. Levinson, S. C. In *Social Intelligence and Interaction* (ed Esther N. Goody) 221–260 (Cambridge University Press, 1995).
2. Tomasello, M. *Origins of Human Communications*. 268 (MIT Press, 2008).
3. Grice, H. P. Meaning. *Philos. Rev.* **66**, 377–388 (1957).
4. Clark, H. H. *Using Language*. (Cambridge University Press, 1996).

5. Bates, E., Benigni, L., Bretherton, I., Camaioni, L. & Volterra, V. *The Emergence of Symbols: Cognition and Communication in Infancy*. (Academic Press, 1979).
6. Johansson, S. *Origins of Language: Constraints on Hypotheses*. (John Benjamins Publishing Company, 2005).
7. Hewes, G. W. Primate communication and the gestural origin of language. *Curr. Anthropol.* **12**, 5–24 (1973).
8. Call, J. & Tomasello, M. *The gestural communication of monkeys and apes* (Lawrence Erlbaum Associates, Mahwah, New York, 2007).
9. Hobaiter, C. & Byrne, R. W. The gestural repertoire of the wild chimpanzee. *Anim. Cogn.* **14**, 747–767 (2011).
10. Pollick, A. S. & de Waal, F. B. M. Ape gestures and language evolution. *Proc. Natl. Acad. Sci.* **104**, 8184–8189 (2007).
11. Hare, B., Melis, A. P., Woods, V., Hastings, S. & Wrangham, R. Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Curr. Biol.* **17**, 619–623 (2007).
12. Melis, A. P., Hare, B. & Tomasello, M. Chimpanzees recruit the best collaborators. *Science* **311**, 1297–1301 (2006).
13. Mitani, J. C. Cooperation and competition in chimpanzees: Current understanding and future challenges. *Evol. Anthropol.* **18**, 215–227 (2009).
14. Wilkinson, R., Leudar, I. & Pika, S. In *Developments in Primate Gesture Research* (eds Simone Pika & Katja Liebal) 199–221 (John Benjamins Publishing Company, 2012).
15. Rossano, F. Sequence organization and timing of bonobo mother-infant interactions. *Interact. Stud.* **14**, 160–189 (2013).
16. Sacks, H., Schegloff, E. A. & Jefferson, G. A simplest systematics for the organization of turn-taking in conversation. *Lang.* **50**, 696–735 (1974).
17. Schegloff, E. A. *Sequence organization in interaction: Volume 1: A primer in conversation analysis*. Vol. 1 (Cambridge University Press, 2007).
18. Levinson, S. C. Recursion in pragmatics. *Lang.* **89**, 149–162 (2013).
19. Stivers, T. *et al.* Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. USA* **106**, 10587–10592 (2009).
20. Indefrey, P. & Levelt, W. J. The spatial and temporal signatures of word production components. *Cognition* **92**, 101–144 (2004).
21. Savage-Rumbaugh, E. S., Brakke, K. E. & Hutchins, S. In *Topics in Primatology Vol. 1: Human Origins Vol. 1 Topics Primatol.* (eds Toshisada Nishida *et al.*) 51–66 (University of Tokyo Press, 1992).
22. Herrmann, E., Hare, B., Call, J. & Tomasello, M. Differences in the cognitive skills of bonobos and chimpanzees. *Plos ONE* **5**, e12438 (2010).
23. Stimpson, C. D. *et al.* Differential serotonergic innervation of the amygdala in bonobos and chimpanzees. *Soc. Cogn. Affect. Neurosci.* **11**, 413–422 (2015).
24. Rilling, J. K. *et al.* Differences between chimpanzees and bonobos in neural systems supporting social cognition. *Soc. Cogn. Affect. Neurosci.* **7**, 369–379 (2012).
25. Bratman, M. E. Shared cooperative activity. *Philosoph. Rev.* **101**, 327–341 (1992).
26. Tomasello, M. & Carpenter, M. Shared intentionality. *Develop. Sci.* **10**, 121–125 (2007).
27. Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. Understanding and sharing intentions: The origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–735 (2005).
28. Boesch, C., Hohmann, G. & Marchant, L. F. *Behavioural Diversity in Chimpanzees and Bonobos* (Cambridge University Press, Cambridge, 2002).
29. Boesch, C. What makes us human (*Homo sapiens*)? The challenge of cognitive cross-species comparison. *J. Comp. Psychol.* **121**, 227–240 (2007).
30. Fröhlich, M., Wittig, R. M. & Pika, S. Should I stay or should I go? Initiation of joint travel in mother–infant dyads of two chimpanzee communities in the wild. *Anim. Cogn.* **19**, 483–500 (2016).
31. Levinson, S. C. & Holler, J. The origin of human multi-modal communication. *Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci.* **369**, 20130302 (2014).
32. Baayen, R. H. *Analyzing linguistic data*. (Cambridge University Press, 2008).
33. Golinkoff, R. M. 'I beg your pardon?': The preverbal negotiation of failed messages. *J. Child. Lang.* **13**, 455–476 (1986).
34. Cheney, D. L., Seyfarth, R. M. & Silk, J. B. The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Anim. Behav.* **50**, 249–257 (1995).
35. Beckers, G. J. & Gahr, M. Neural processing of short-term recurrence in songbird vocal communication. *Plos One* **5**, e11129 (2010).
36. Hare, B. Can competitive paradigms increase the validity of experiments on primate social cognition? *Anim. Cogn.* **4**, 269–280 (2001).
37. Marshall-Ball, L. & Slater, P. J. B. Duet singing and repertoire use in threat signalling of individuals and pairs. *Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci.* **271**, 440–443 (2004).
38. Malenky, R. K. & Wrangham, R. W. A quantitative comparison of terrestrial herbaceous food consumption by *Pan paniscus* in the Lomako Forest, Zaire, and *Pan troglodytes* in the Kibale Forest, Uganda. *Am. J. Primatol.* **19**, 999–1011 (1994).
39. Boesch, C. & Boesch-Achermann, H. *The Chimpanzees of the Tai Forest: Behavioural Ecology and Evolution*. (Oxford University Press, 2000).
40. Kuroda, S. In *Understanding chimpanzees* (eds P. G. Heltne & L. A. Marquardt) 184–193 (Harvard University Press, 1989).
41. De Lathouwers, M. & Van Elsacker, L. Comparing infant and juvenile behavior in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): A preliminary study. *Primates* **47**, 51–55 (2006).
42. Pika, S. In *The Social Origins of Language: Early Society, Communication and Polymodality* (eds D. Dor, C. Knight & J. Lewis) 129–140 (Oxford University Press, 2014).
43. De Waal, F. B. Primates—a natural heritage of conflict resolution. *Science* **289**, 586–590 (2000).
44. King, B. J. *The dynamic dance: Nonvocal communication in African great apes*. (Harvard University Press, 2004).
45. Surbeck, M. & Hohmann, G. Intersexual dominance relationships and the influence of leverage on the outcome of conflicts in wild bonobos (*Pan paniscus*). *Behav. Ecol. Sociobiol.* **67**, 1767–1780 (2013).
46. Nishida, T. & Hiraiwa-Hasegawa, M. In *Primate Societies* (eds B. B. Smuts *et al.*) 165–177 (The University of Chicago Press, 1987).
47. Watts, D. P., Mitani, J. C. & Sherrow, H. M. New cases of inter-community infanticide by male chimpanzees at Ngogo, Kibale National Park, Uganda. *Primates* **43**, 263–270 (2002).
48. Furuichi, T. Female contributions to the peaceful nature of bonobo society. *Evol. Anthropol.* **20**, 131–142 (2011).
49. Idani, G. Social relationships between immigrant and resident bonobo (*Pan paniscus*) females at Wamba. *Folia Primatol.* **57**, 83–95 (1991).
50. Hohmann, G. & Fruth, B. In *Among African Apes: Stories and Photos from the Field* (eds Martha M. Robbins & Christophe Boesch) 61–76 (University of California Press, 2011).
51. Hohmann, G. & Fruth, B. Intra- and inter-sexual aggression by bonobos in the context of mating. *Behaviour* **140**, 1389–1413 (2003).
52. Freese, J. L. & Amaral, D. G. In *The Human Amygdala* (eds P. J. Whalen & E. A. Phelps) 3–42 (The Guilford Press, 2009).
53. Herrmann, E., Hare, B., Cissewski, J. & Tomasello, M. A comparison of temperament in nonhuman apes and human infants. *Develop. Sci.* **14**, 1393–1405 (2011).
54. Pepperberg, I. M. *The Alex Studies, Cognitive and Communicative Abilities of Grey Parrots*. (Harvard University Press, 1999).
55. Pika, S. & Bugnyar, T. The use of referential gestures in ravens (*Corvus corax*) in the wild. *Nat. Com.* **2**, 1–5 (2011).
56. Hohmann, G. & Fruth, B. Lui Kotal: A new site for field research on bonobos in the Salonga National Park. *Pan Africa News* **10**, 25–27 (2003).
57. Kano, T. *The last ape: Pygmy chimpanzee behavior and ecology*. (Stanford University Press, 1992).

58. Wrangham, R. W., Clark, A. P. & Isabirye-Basuta, G. Female social relationships and social organization of the Kibale Forest chimpanzees. *Topics Primatol.* **1**, 81–98 (1992).
59. Altmann, J. Observational study of behaviour: Sampling methods. *Behaviour* **49**, 227–267 (1974).
60. Nishida, T., Kano, T., Goodall, J., McGrew, W. C. & Nakamura, M. Ethogram and ethnography of Mahale chimpanzees. *Anthropol. Sci.* **107**, 141–188 (1999).
61. Pika, S., Liebal, K. & Tomasello, M. Gestural communication in subadult bonobos (*Pan paniscus*): Gestural repertoire and use. *Am. J. Primatol.* **65**, 39–51 (2005).
62. Pika, S. Gestures of apes and pre-linguistic human children: Similar or different? *First Lang.* **28**, 116–140 (2008).
63. van de Pol, M. & Wright, J. A simple method for distinguishing within-versus between-subject effects using mixed models. *Anim. Behav.* **77**, 753–758 (2009).
64. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
65. R Core Team R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/> (2013).
66. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. *R package version 1*, 1–7 (2014).
67. Dobson, A. J. *An Introduction to Generalized Linear Models*. (Chapman & Hall/CRC, 2002).
68. Schielzeth, H. Simple means to improve the interpretability of regression coefficients. *Meth. Ecol. Evol.* **1**, 103–113 (2010).
69. Quinn, G. P. & Keough, M. J. *Experimental design and data analysis for biologists*. (Cambridge University Press, 2002).

## Acknowledgements

This paper is dedicated to the memory of B. M. Siemers. We are indebted to C. Boesch, G. Hohmann, M. N. Muller and R. W. Wrangham for providing the opportunity to carry out research at *Tai South*, *LuiKotale* and *Kanyawara*. Without their continuous support and constructive criticism this work would not have been possible. We are grateful to the teams of the *LuiKotale* and *Wamba* Bonobo projects and the teams of the *Kibale* and *Tai* Chimpanzee Projects for engaging in invaluable assistance during the fieldwork of this study. We thank the Centre de Recherche en Écologie et Forêt (CREF; DRC), the Institut Congolaise pour la Conservation de la Nature (ICCN; DRC), the Makerere University Biological Field Station (MUBFS; Uganda), the Ministère de l'Enseignement Supérieure et de la Recherche Scientifique (Côte d'Ivoire), the Ministère de Recherche Scientifique (DRC), the Office Ivoirien des Parcs et Réserves (OIPR; Côte d'Ivoire), the Uganda National Council for Science and Technology (UNCST; Uganda) and the Uganda Wildlife Authority (UWA; Uganda) for granting permission to conduct research at Salonga National Park (DRC), the Luo Reserve (DRC), Kibale National Park (Uganda) and Tai National Park (Côte d'Ivoire). We are grateful to K. Graham and L. Alcayna-Stevens for invaluable assistance with data collection, to R. Mundry for statistical advice and to D. Enigk for proofreading. We thank C. Crockford, D. Dechmann and T. Deschner for valuable discussions during the preparation of the manuscript and M. Krug for her steady support. A Dissertation Fieldwork Grant of the Wenner-Gren Foundation to M.F., a Research Grant of the Leakey Foundation to P.K. and a Sofa-Kovalevskaja Award of the Humboldt Foundation to S.P. generously supported the project.

## Author Contributions

S.P. and M.F. designed the project. M.F. and P.K. collected data. M.F. and G.M. coded the data. M.F. analysed the data. M.F. and S.P. wrote the paper. B.F., T.F., R.M.W. and P.K. commented on the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Fröhlich, M. *et al.* Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Sci. Rep.* **6**, 25887; doi: 10.1038/srep25887 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>